

Tematy projektów z Metod Nieparametrycznych  
(na ostatniej stronie znajdują się zasady oceniania)

#### **Zad. 1**

Za pomocą odpowiednich symulacji porównać moc klasycznego testu równości dwóch wartości oczekiwanych i testu Manna-Whitneya. Porównanie przeprowadzić w przypadku, gdy dane pochodzą z dwóch rozkładów normalnych różniących się wartościami oczekiwanyymi. Uzyskane wyniki (np. krzywe mocy testu) należy przedstawić na odpowiednich wykresach ilustrujących zmiany mocy testów w zależności od:

- liczby danych,
- różnicy (ilorazu) odchyleń standardowych.

#### **Zad. 2**

Za pomocą odpowiednich symulacji zbadać odsetek odrzuceń prawdziwej hipotezy głównej (tzn. błąd I rodzaju) w klasycznym teście równości dwóch wartości oczekiwanych i w teście Manna-Whitneya w przypadku, gdy dane w obu próbkach pochodzą z rozkładu normalnego o takiej samej wartości oczekiwanej, lecz różnią się odchyleniami standardowymi. W każdym przypadku wyznaczyć empiryczne wartości krytyczne. Porównać te wyniki z rezultatami zastosowania statystyki  $t$  (np. funkcja t.test)

Uzyskane wyniki należy opisać zależności od:

- liczby danych w każdej z próbek,
- różnicy odchyleń standardowych.

#### **Zad. 3**

Za pomocą odpowiednich symulacji zbadać odsetek odrzuceń hipotezy głównej w teście Manna-Whitneya i w teście równości dwóch wartości oczekiwanych (funkcja t.test) w przypadku, gdy jedno dane pochodzą z rozkładu normalnego, a drugie np. z rozkładu  $\chi^2$  (który nie jest rozkładem symetrycznym). W trakcie symulowania danych należy pamiętać, by dane spełniały hipotezę główną (najlepiej je zestandaryzować).

Uzyskane wyniki należy opisać zależności od:

- liczby danych,
- liczby stopni swobody generowanego rozkładu  $\chi^2$ .

#### **Zad. 4**

Za pomocą odpowiednich symulacji zbadać odsetek odrzuceń prawdziwej hipotezy głównej (tzn. błąd I rodzaju) w teście Manna-Whitneya i w teście równości dwóch wartości oczekiwanych (funkcja t.test) w przypadku, gdy dane pochodzą z rozkładów  $\chi^2$  o różnej liczbie stopni swobody. W trakcie symulowania danych należy pamiętać, by dane spełniały hipotezę główną (tzn. by miały tę samą wartość oczekiwaną. Zbadać wpływ standaryzacji danych na uzyskane wyniki.

Uzyskane wyniki należy opisać zależności od:

- liczby danych,
- liczby stopni swobody generowanych rozkładów  $\chi^2$ .

#### **Zad. 5**

Za pomocą odpowiednich symulacji zbadać odsetek odrzuceń prawdziwej hipotezy głównej (tzn. błąd I rodzaju) w teście istotności współczynnika korelacji w przypadku, gdy dane pochodzą z rozkładu normalnego o różnych parametrach. W trakcie symulowania danych należy pamiętać, by dane spełniały hipotezę główną. Porównać uzyskane wyniki dla wsp. korelacji Pearsona i Spearmana.

Uzyskane wyniki należy opisać w zależności od:

- liczby danych,
- parametrów rozkładów.

#### **Zad. 6**

Za pomocą odpowiednich symulacji zbadać odsetek odrzuceń prawdziwej hipotezy głównej (tzn. błąd I rodzaju) w teście istotności współczynnika korelacji w przypadku, gdy dane pochodzą z rozkładów  $\chi^2$  (o różnych stopniach swobody). W trakcie symulowania danych należy pamiętać, by dane spełniały hipotezę główną. Porównać uzyskane wyniki dla wsp. korelacji Pearsona i Spearmana.

Uzyskane wyniki należy opisać w zależności od:

- liczby danych,
- liczby stopni swobody generowanych rozkładów  $\chi^2$ .

#### **Zad. 7**

Za pomocą odpowiednich symulacji zbadać odsetek odrzuceń prawdziwej hipotezy głównej (tzn. błąd I rodzaju) w teście istotności współczynnika korelacji w przypadku, gdy dane pochodzą z rozkładu o grubych ogonach (np. *t*-Studenta o różnych stopniach swobody). W trakcie symulowania danych należy pamiętać, by dane spełniały hipotezę główną. Porównać uzyskane wyniki dla wsp. korelacji Pearsona i Spearmana.

Uzyskane wyniki należy opisać w zależności od:

- liczby danych,
- liczby stopni swobody generowanych rozkładów  $\chi^2$ .

#### **Zad. 8**

Za pomocą odpowiednich symulacji zbadać odsetek odrzuceń prawdziwej hipotezy głównej (tzn. błąd I rodzaju) w teście istotności współczynnika korelacji w przypadku, gdy dane pochodzą z rozkładu  $\chi^2$  i rozkładu  $-1 \cdot \chi^2$  (tzn. generujemy dane z rozkładu  $\chi^2$  i mnożymy je przez -1). W trakcie symulowania danych należy pamiętać, by dane spełniały hipotezę główną. Porównać uzyskane wyniki dla wsp. korelacji Pearsona i Spearmana.

Uzyskane wyniki należy opisać w zależności od:

- liczby danych,
- liczby stopni swobody generowanego rozkładu  $\chi^2$ .

#### **Zad. 9**

Za pomocą odpowiednich symulacji zbadać odsetek odrzuceń prawdziwej hipotezy głównej (tzn. błąd I rodzaju) w ANOVA i w teście Kruskala-Wallisa w przypadku, gdy generowane dane nie pochodzą z rozkładu normalnego (ale spełniają pozostała założenia ANOVA). W tym celu można wygenerować trzy próbki danych z tego samego rozkładu t-Studenta.

Uzyskane wyniki należy opisać w zależności od:

- liczebności generowanych próbek,
- liczby stopni swobody generowanego rozkładu t-Studenta (proszę pamiętać, że od stopni swobody zależy odchylenie standardowe).

#### **Zad. 10**

Za pomocą odpowiednich symulacji zbadać odsetek odrzuceń prawdziwej hipotezy głównej (tzn. błąd I rodzaju) w ANOVA i teście Kruskala-Wallisa w przypadku, gdy generowane dane mają różne wariancje w grupach. W tym celu można wygenerować trzy próbki danych z rozkładu normalnego: dwie o tym samym odchyleniu stand., a trzecia o innym odchyleniu stand.

Uzyskane wyniki należy opisać w zależności od:

- liczebności generowanych próbek,
- różnicy (ilorazu) odchyleń standardowych.

#### **Zad. 11**

Za pomocą odpowiednich symulacji zbadać odsetek odrzuceń prawdziwej hipotezy głównej (tzn. błąd I rodzaju) w ANOVA i teście Kruskala-Wallisa w przypadku gdy podgrupy mają różną liczebność.

Rozważyć przypadek trzech grup.

Uzyskane wyniki należy opisać w zależności od:

- liczebności generowanych grup i różnic pomiędzy liczebnościami poszczególnych grup.

#### **Zad. 12**

Za pomocą odpowiednich symulacji zbadać moce ANOVA i testu Kruskala-Wallisa w zależności od liczebności grup. Rozważyć różne formy postawienia fałszywej hipotezy głównej.

Uzyskane wyniki należy opisać w zależności od:

- liczebności generowanych grup,
- postaci fałszywej hipotezy głównej.

#### **Zasady oceniania**

Projekt należy wykonać w programie R. Wykonanie projektu obejmuje: opis (w języku polskim ) przeprowadzonych badań wraz z wynikającymi z nich wnioskami oraz skrypt w R użyty do przeprowadzenia potrzebnych obliczeń i utworzenia wykresów.